



## The use of Softwares in Business Research

**Dr. Satyabrata Dash**  
**Principal**  
**SMIT- PGCMS**



## *Statistical software*

- Statistical software are specialized computer programs for statistical analysis.
  - Open source statistical packages
  - Public domain statistical packages
  - Freeware statistical packages
  - Proprietary statistical packages



## *Statistical software*

- **Open source statistical packages**
  - **ADMB** : a software for non-linear statistical modeling based on C++
  - **DAP** : A free replacement for SAS
  - **Fityk** : nonlinear regression software
  - **OpenEpi** : A web-based, open source, operating-independent series of programs for use in epidemiology and statistics
  - **SciPy (statsmodels)** : regression, plotting, example datasets, GLM, time series analysis, Non-parametric statistics, ANOVA
  - **PSPP** : A free software alternative to IBM SPSS Statistics
  - **R** : A free implementation of the S language

# *Statistical software*

- **Public domain statistical packages**
  - **CSPRO**
    - Developed by US census bureau and ICF International
    - Used for entering, editing, tabulating, mapping and disseminating census and survey data.
  - **Epi Info**
    - public domain statistical software for epidemiology developed by Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia (USA).
    - The program allows for electronic survey creation, data entry, and analysis. Within the analysis module, analytic routines include t-tests, ANOVA, nonparametric statistics, cross tabulations and stratification with estimates of odds ratios, risk ratios, and risk differences, logistic regression (conditional and unconditional), survival analysis (Kaplan Meier and Cox proportional hazard), and analysis of complex survey data.
  - **X-12-ARIMA**
    - Developed by U.S. Census Bureau's software package for seasonal adjustment.

## *Statistical software*

- Freeware statistical packages
  - WinBUGS : Bayesian analysis using Markov chain Monte Carlo methods
  - Winpepi : package of statistical programs for epidemiologists

# *Statistical software*

- **Proprietary statistical packages**
  - **GraphPad InStat** : Very simple with lots of guidance and explanations
  - **GraphPad Prism** : Biostatistics and nonlinear regression with clear explanations
  - **IBM SPSS Statistics** : comprehensive statistics package
  - **IBM SPSS Modeler** : comprehensive data mining and text analytics workbench
  - **MATLAB** : programming language with statistical features
  - **SAS** : comprehensive statistical package
  - **SPSS** - : Statistical Package for the Social Sciences
  - **StatsDirect** : statistics package designed for biomedical, public health and general health science uses



## *Statistical software*

### *Microsoft Excel add-ons*

- **Analyse-it** : add-on to Microsoft Excel for statistical analysis
- **NumXL** : add-on to Microsoft Excel for general statistics and Econometrics
- **RegressIt** : add-on to Microsoft Excel for multivariate data analysis and linear regression (freeware)
- **SigmaXL** : add-on to Microsoft Excel for statistical and graphical analysis
- **SPC XL** : add-on to Microsoft Excel for general statistics
- **Stats Helper** : add-on to Microsoft Excel for descriptive statistics and Six Sigma

## *Statistical software*

- Most commonly used software in the field of Pharmacy
  - SAS
  - SPSS
  - GraphPad InStat
  - GraphPad Prism



# *Statistical software - SPSS*

- **SPSS: Statistical Package for the Social Sciences**
- **Current version: IBM SPSS statistics (2016); latest version 24**
- **Methods:**
  - **Descriptive statistics: Cross tabulation, Frequencies, Descriptives, Explore, Descriptive Ratio Statistics**
  - **Bivariate statistics: Means, t-test, ANOVA, Correlation (bivariate, partial, distances), Nonparametric tests**
  - **Prediction for numerical outcomes: Linear regression**
  - **Prediction for identifying groups: Factor analysis, cluster analysis (two-step, K-means, hierarchical), Discriminant**
- **SPSS statistics desktop platforms:**
  - **Linux, Mac OS, Windows**

## About SPSS Inc.

- SPSS Inc. is a leading worldwide provider of predictive analytics software and solutions.
- Founded in 1968, today SPSS has more than 250,000 customers worldwide, served by more than 1,200 employees in 60 countries.

# What is SPSS?

- Statistical Package for the Social Sciences
- Superior Performing Statistical Software
- Statistical Products and Software Solutions
  
- Processing Questionnaires
- Reporting in Tables and Graphs
- Analyzing: Means, Chi-square, Regression, ...



# Research Process

- Problem definition
- Research objectives
- Desk Research
- Field Research
  - Qualitative
  - Quantitative: constructing a questionnaire
- Collecting and Analyzing data
- Writing and Presenting the final research report

# Translate the Questionnaire into SPSS

Questions in the questionnaire  
are mapped into  
Variables in SPSS

# Measurement in Market Research

- Question response formats
- Scale characteristics
- Levels of measurement



## Question-response formats

- Closed-Ended
- Open-Ended with numerical response
- Open-Ended with text response
- Multiple response questions

## Respons-format :: Closed-Ended (1)

5.How is your satisfaction with the customer service of the staff of Suxes?

- ☐ Excellent
- ☐ Good
- ☐ Bad
- ☐ Very bad

## Coding the answers

1 = Excellent

2 = Good

3 = Bad

4 = Very bad



## Respons-format :: Closed-Ended (2)

11. Please indicate your gender.

☐ Female

☐ Male

Codes:

1 = Female

2 = Male

Open-ended with numerical response

2. What is your average expenditure in the restaurant on a weekly basis?

..... euro per week

10. For how many years have you been registered as a student at Pandion University?

..... year(s)

## Open-ended with text response

1. I would like to have the assortment extended with the following products:

.....

### Processed by

- Coding manually afterwards or
- Typing the answers literally (text variable)  
(Auto recode afterwards)



# Scale characteristics

- Description
- Order
- Distance
- Nominal
- Ordinal
- Scale

Levels of Measurement

# Levels of measurement

## Nominal

- Your course at Pandion  
(Marketing, Int. Business and languages, Int. Business administration, Management and law, Health studies, Security);
- Smoker or non smoker (yes, no);
- Choice of a supermarket  
(A&P, Wal-Mart, Sears, Aldi, other).

## Levels of measurement

### Ordinal

In your opinion, would you say the prices at Wal-Mart are

- ☐ Higher than Sears
- ☐ About the same as Sears
- ☐ Lower than Sears

# Levels of measurement

## Ordinal

What is your age?

- ☐ 15–<25
- ☐ 25–<40
- ☐ 40–<60
- ☐ 60–<90

# Levels of measurement

## Ordinal

Statement	Strongly disagree				Strongly agree
a. I always look for bargains	1	2	3	4	5
b. I enjoy being outdoors	1	2	3	4	5
c. I love to cook	1	2	3	4	5



## Levels of measurement

### Scale

Please indicate your age.

\_\_\_\_ Years

How much do you think a typical purchaser of a \$100,000 term life insurance policy pays per year for that policy?

\$ \_\_\_\_\_

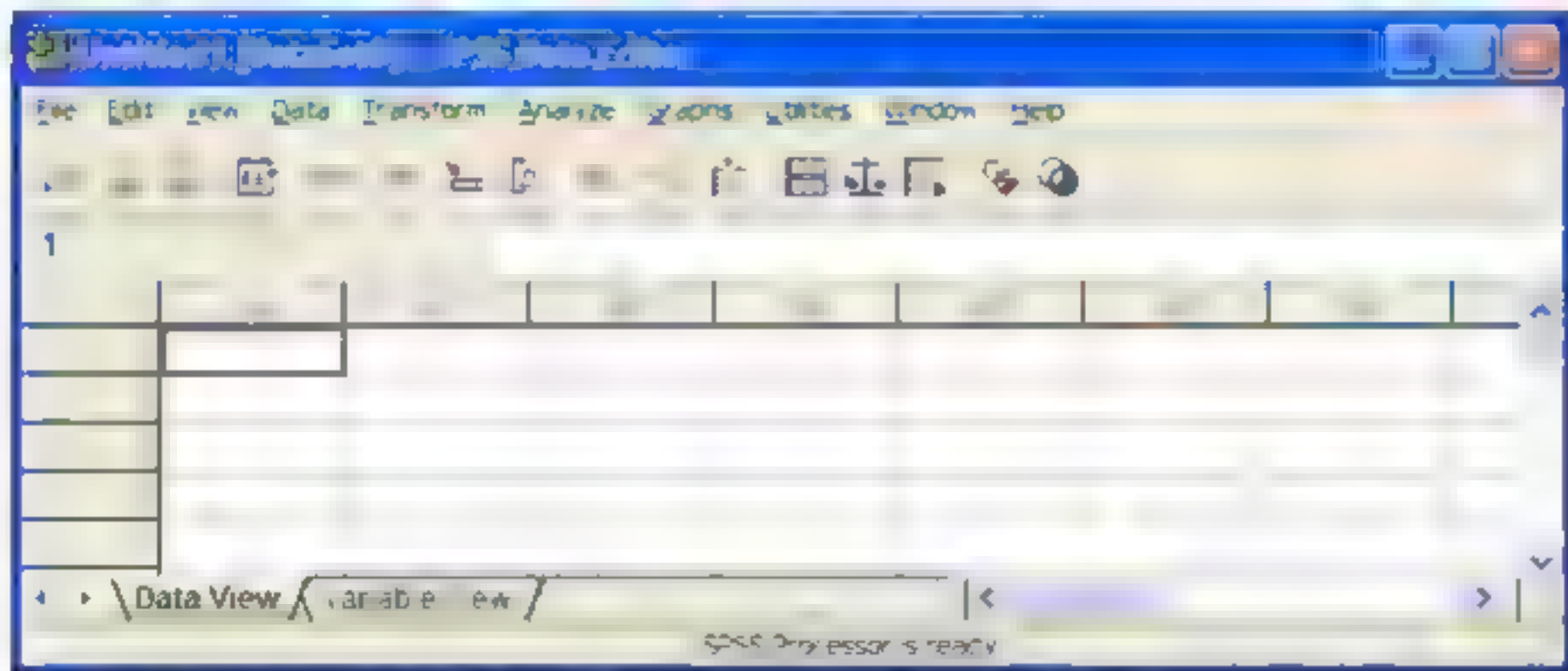
# Coding data and the data code book

## Questions → Variables

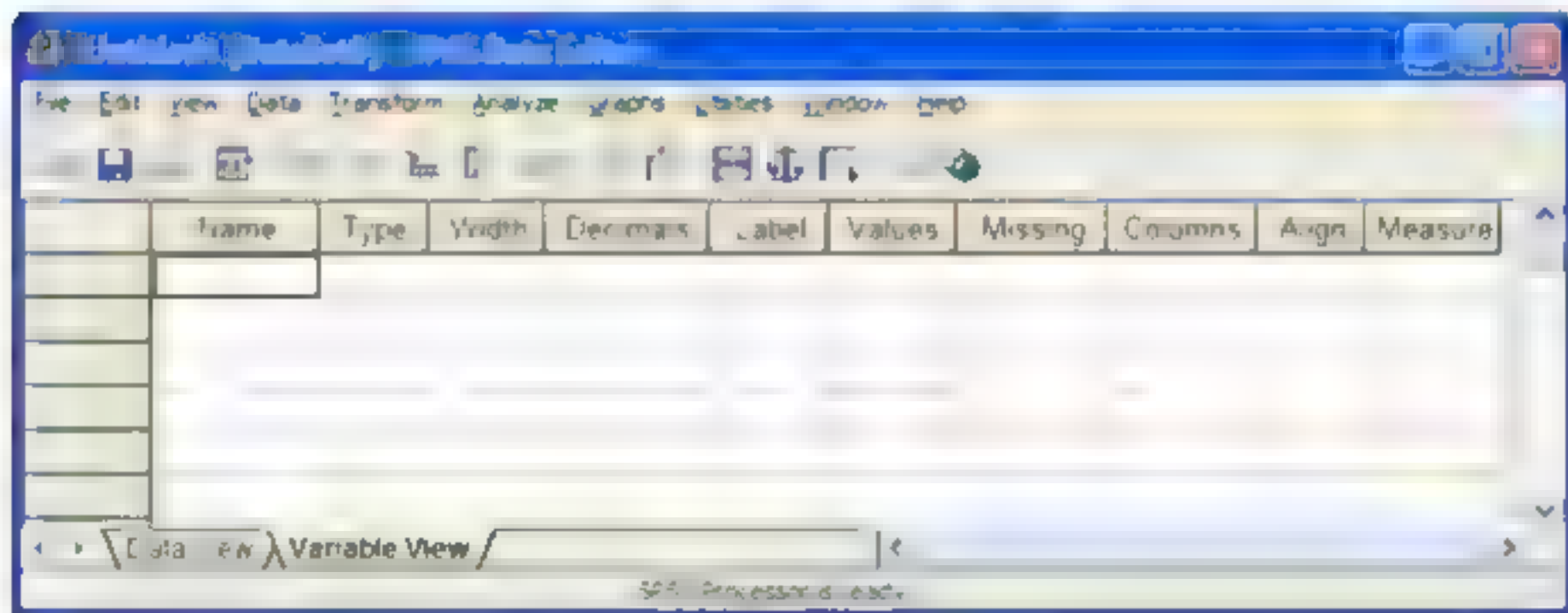
- Name of the variable
- Variable label
- Value labels (data codes)
- Level of measurement (Measure)

# The SPSS Data Editor

## Data View



## The SPSS Data Editor Variable View



# The SPSS Data Editor

## Variable view

- Name
- Type (Numeric)
- Label
- Values (= the codes of the answers)
- Measure (= Level of Measurement)



# SPSS Menu's

- Analyze
  - Frequencies
  - Cross tabs
  - Tables
  - 
  - 
  -

# SPSS Menu's

- Graphs
  - Bar
  - Pie
  - Histogram
  - Line
  - Boxplot

# SPSS Output

- Separate file in Output Viewer
- Inline Editing of Tables
- Chart Editor for Graphs

Don't forget to save

- Data file
- Output file



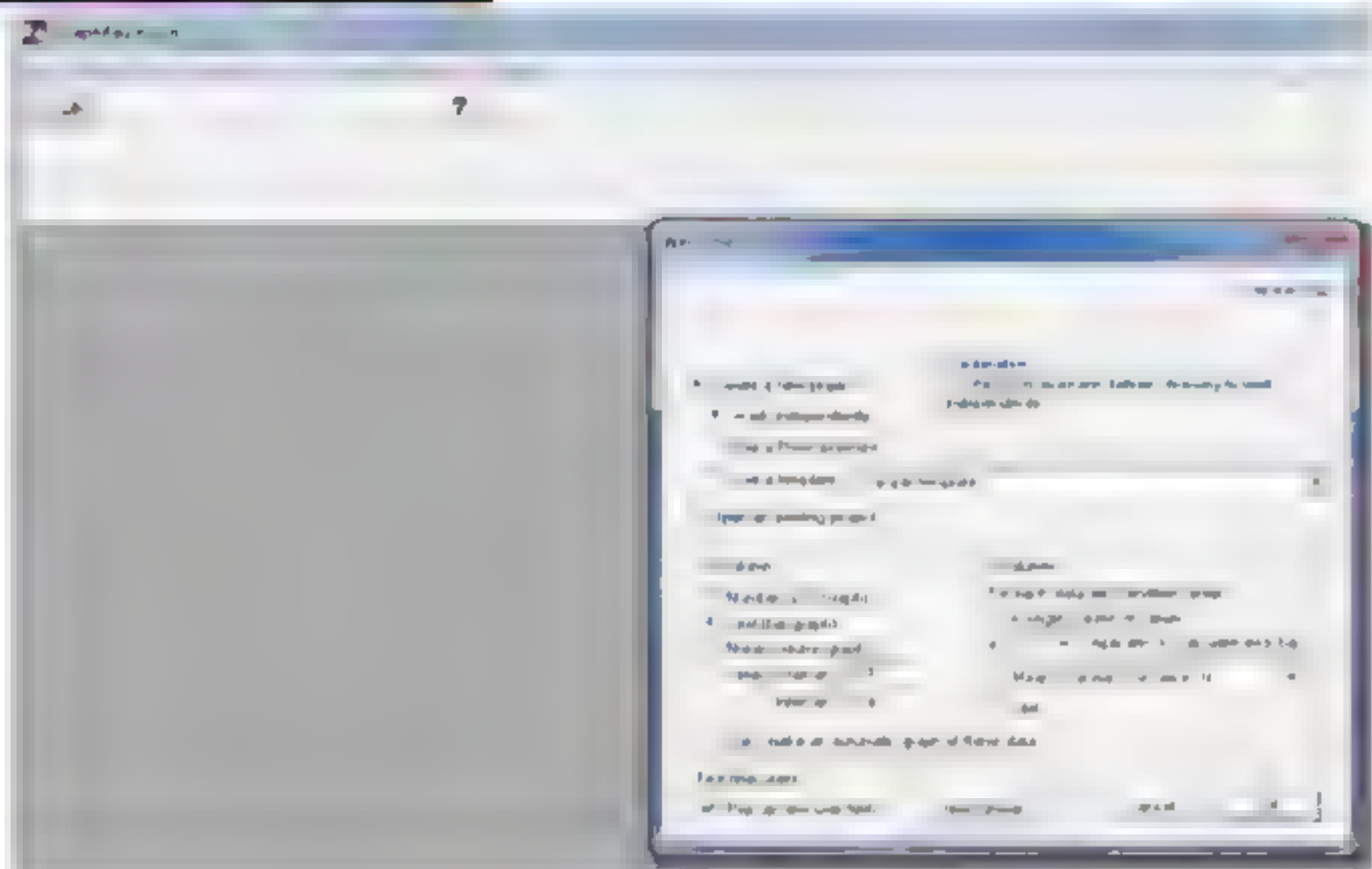
## *Statistical software – SAS (Statistical Analysis System)*

- Developed by: North Carolina State University in 1966
- SAS provides a graphical point-and-click user interface for non-technical users and more advanced options through the SAS programming language.
- The DATA step has two phases, compilation and execution.
- The SAS software suite has more than 200 components, including
  - Base SAS - Basic procedures and data management
  - SAS/STAT - Statistical analysis
  - SAS/GRAPH - Graphics and presentation
  - SAS/OR - Operations research
  - SAS/ETS - Econometrics and Time Series Analysis
  - SAS/QC - Quality control SAS/INSIGHT - Data mining
  - SAS/PH - Clinical trial analysis





# Statistical software – GraphPad Prism





# Parametric vs. non-parametric tests

Normal	Any
Homogeneous	Any
Ratio or Interval	Ordinal or Nominal
Independent	Any
Mean	Median
Can draw more conclusions	Simplicity, Less affected by outliers

Choosing parametric test	Choosing a non-parametric test
Pearson	Spearman
Independent-measures t-test	Mann-Whitney test
One-way, independent-measures ANOVA	Kruskal-Wallis test
Matched-pair t-test	Wilcoxon test
One-way, repeated measures ANOVA	Friedman's test

## *software for parametric analysis*

- GraphPad InStat (not having analytical futures for analysis of Two-way ANOVA)
- GraphPad Prism
- IBM SPSS Statistics
- SAS
- StatsDirect
- MATLAB

## *software for non-parametric analysis*

- GraphPad InStat
- GraphPad Prism
- IBM SPSS Statistics
- SAS
- StatXact
- SURVSOFT
- R Software

## *software for non-parametric analysis*

- **Models:** descriptive statistics, statistical models, inference and statistical tests.
- **Methods:**
  - Kruskal-Wallis one-way analysis of variance
  - Mann-Whitney U
  - Siegel-Tukey test
  - Spearman's rank correlation coefficient
  - Wilcoxon signed-rank test
  - Kaplan-Meier
  - Friedman two-way analysis of variance
  - McNemar's test
  - sign test

*Thank you*